Good potentials for almost isomorphism of countable state Markov shifts

Jerome Buzzi, Mike Boyle, Ricardo Gomez

GOOD POTENTIALS FOR ALMOST ISOMORPHISM OF COUNTABLE STATE MARKOV SHIFTS

MIKE BOYLE, JEROME BUZZI, AND RICARDO GÓMEZ

Abstract. Almost isomorphism is an equivalence relation on countable state Markov shifts which provides a strong version of Borel conjugacy; still, for mixing SPR shifts, entropy is a complete invariant of almost isomorphism [2]. In this paper, we establish a class of potentials on countable state Markov shifts whose thermodynamic formalism is respected by almost isomorphism.

Contents

1. Introduction

The classical subshifts of finite type [1, 10, 12] admit the following natural generalization. A *countable state Markov shift* is (X, S) where $X \subset V^{\mathbb{Z}}$ for some countable (maybe finite) set V and for some $E \subset V^2$:

$$X = \{x \in V^{\mathbb{Z}} : \forall n \in \mathbb{Z}\ (x_n, x_{n+1}) \in E\}$$

and $S : X \to X$ defined by

$$S((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}.$$

In other words, X is the set of bi-infinite paths on the directed graph (V, E) together with the left-shift S. The one-sided version (X_+, S_+) is given by $X_+ = \{x_0x_1 \cdots \in V^{\mathbb{N}} : x \in X\}$ and $S_+((x_n)_{n\in\mathbb{N}}) = (x_{n+1})_{n\in\mathbb{N}}$. (We shall sometimes use the same letter for both the map and the space.) A countable state Markov shift is *transitive*, or *irreducible*, if for any vertices u, v in the underlying directed graph there is a path from u to v; and it is *mixing* if for given $(u, v) \in V^2$, for all but finitely many k, this path may be chosen to have length k. The *topological entropy*, or *Gurevič entropy*

2000 *Mathematics Subject Classification.* Primary: 37B10; Secondary: 37B40, 37C99, 37D35.

Key words and phrases. entropy; countable state Markov shift; topological Markov chain; almost isomorphism; entropy conjugacy; magic word; strong positive recurrence; thermodynamic formalism; variational principle; equilibrium state; Artin-Mazur zeta function; smooth ergodic theory.

This work was partly supported by NSF Grant 0400493 (Boyle).

[7], of (X,S) is $h(S) = \sup_\mu h(S,\mu)$, the supremum of the entropies of S-invariant probability measures (see [24] for background on entropy).

STANDING CONVENTION: for the rest of this paper, unless there is an explicit qualification, "Markov shift" means an irreducible countable (maybe finite) state Markov shift of finite topological entropy.

A Markov shift (X,S) is *strongly positive recurrent* or SPR [22, 8, 6, 17, 16, 3] if it admits an invariant probability measure μ with $h(S,\mu) = h(S)$ and

$$\limsup_{n\to\infty} (1/n) \log \mu \left(X \setminus \bigcup_{k=0}^{n} S^{-k} U \right) < 0$$

for all nonempty open $U \subset X$. In the terminology of [9], the SPR shifts are the positive recurrent Markov shifts whose defining directed graphs have adjacency matrices which are *stable positive.* There are other characterizations of SPR Markov shifts (assembled in [2, Prop. 2.3]). The SPR shifts are the natural large class of Markov shifts which still retain certain key properties of shifts of finite type.

Almost isomorphism of Markov shifts yields a strong version of Borel conjugacy (recalled in Section 2). In particular, almost isomorphism of Markov shifts induces a bijection between their sets of ergodic fully supported measures, simultaneously defining isomorphisms of all the corresponding measurable systems. This gives an identification of the measures of maximal entropy, when they exist, by a map for which the coding time (in the SPR case) has an exponential tail. Our main result in [2] shows that for mixing SPR Markov shifts, topological entropy is a complete invariant of almost isomorphism.

Thermodynamic formalism generalizes the notion of measure of maximal entropy to equilibrium measure (2.1) of a function (potential). An equilibrium measure for a "nice" potential should be fully supported and ergodic (remember that we assume our systems to be irreducible). Almost isomorphism respects the class of such measures, so this begs the question, is there a class of reasonably nice potentials which together with their equilibrium measures are respected by almost isomorphism? The usual classes of nice potentials (locally Hölder continuous or with summable variations) are not preserved by almost isomorphism.

In this paper we introduce a new class of potentials called *relatively regular* which, on the one hand, are nice enough to guarantee existence of a unique equilibrium measure which is fully supported and, on the other hand, are invariant under almost isomorphism. This class contains the positive recurrent potentials of summable variation, but necessarily contains certain less regular potentials as well (though it does not include all potentials which even smooth functions can generate by coding when there is non-uniform expansion). Our main result is that almost isomorphisms leave globally invariant the thermodynamic formalism of the bounded relatively regular functions.

In Sec 2, we provide definitions and state our main result, Theorem 2.6: the thermodynamic formalism of the bounded relatively regular functions is respected by almost isomorphism. The uniqueness of their equilibrium measures is established (in greater generality) in Section 3; the existence is established (in greater generality) in 4; and the correspondence under almost isomorphism is established in Section 5.

2. Statement of result

Given a Borel map $S : X \to X$, a real-valued Borel measurable function f from X to $\mathbb{R}$, and an S-invariant Borel probability μ, we define

$$P(S, f, \mu) = h_\mu(S) + \int f \; d\mu \ , \qquad \text{and}$$
$$P(S, f) = \sup_\mu P(S, f, \mu) \ .$$

We take $P(S, f)$ as our definition of the *pressure* of (S, f). $P(S, f, \mu)$ might not be defined for some μ; the supremum defining $P(S, f)$ is taken over the well defined $P(S, f, \mu)$.

Definition 2.1. An *equilibrium measure* of f (also called an equilibrium state) is an S-invariant Borel probability μ such that $P(S, f, \mu) = P(S, f)$.

For background see [24, 9, 5, 14, 18].

Next we recall the definition of almost isomorphism from [2]. A map $\varphi : S \to T$ between Markov shifts is a *one-block code* if there is a function Φ from the symbol set of S into the symbol set of T such that $(\varphi x)_n = \Phi(x_n)$, for all x and n (note that $\varphi x = \varphi(x)$). A T-word W (of length $|W|$) is a *magic word* for such a map φ if the following hold. We denote by $x[a, b]$ the restriction of the sequence x to the indices $i = a, a + 1, \ldots, b$.

(1) If $y \in T$ and $\{n \in \mathbb{Z} : y[n, n + |W| - 1] = W\}$ is unbounded above and unbounded below, then y has a preimage under φ.
(2) There is an integer I such that whenever C is a T-word and two points x and x' of S satisfy $(\varphi x)[0, 2|W|+|C|-1] = WCW = (\varphi x')[0, 2|W|+|C|-1]$, then $x[I, I + |W| + |C| - 1] = x'[I, I + |W| + |C| - 1]$.

(In constructions, the integer I of the last condition can generally be chosen to be zero.) It follows from (2) that the preimage in (1) is unique.

Definition 2.2. Markov shifts S and T are *almost isomorphic* if there exist a Markov shift R and one-block codes $R \to S$, $R \to T$ each of which is injective with a magic word. Such a pair of maps is an *almost isomorphism* of S and T.

Our interest in almost isomorphism is largely explained by the following proposition, copied from [2, Proposition 3.4]. (We use "vertex shifts" in this paper rather than the "edge shifts" of [2], but this is only a matter of notation.)

Proposition 2.3. *Suppose S and T are almost isomorphic Markov shifts. Then $h(S) = h(T)$, and there are Borel subsets K and K' of S and T, collections of invariant probability measures $\mathcal{M}(K), \mathcal{M}(K')$ on K and K' and a shift-commuting Borel-measurable bijection $\gamma : K \to K'$, such that the following hold.*

(1) *K and K' are residual subsets of S and T (contain dense G_δ sets).*
(2) *The map γ induces a bijection $\mathcal{M}(K) \to \mathcal{M}(K')$ ($\mu \mapsto \mu'$, say) such that for each such pair μ, μ' the map γ induces an isomorphism $\gamma : (S, \mu) \to (T, \mu')$, which is a magic word isomorphism when μ and μ' have full support.*
(3) *$\mathcal{M}(K)$ and $\mathcal{M}(K')$ contain all ergodic shift-invariant Borel probabilities on S and T with full support, and these correspond under γ.*
(4) *If S is SPR, then so is T, and γ is an entropy-conjugacy from S to T.*

In the proposition, "full support" means that the measure is nonvanishing on all nonempty open sets, and "entropy-conjugacy" means that there exists $\epsilon > 0$ such that the sets $\mathcal{M}(K)$, $\mathcal{M}(K')$ contain all invariant ergodic Borel probabilities with entropy greater than $h(S) - \epsilon$. We denote by $\mathcal{M}^{\text{erg}}_{\text{supp}}(S)$ the set of ergodic fully supported S-invariant Borel probabilities.

Given a Markov shift S and a nonempty collection $\mathcal{W}$ of S-words (words appearing in points of S), let $S_{\mathcal{W}}$ denote the set of all points x in S which see $\mathcal{W}$ words infinitely often in the past and future[1]. (When $\mathcal{W} = \{W\}$ we may use the notation S_W.) If W is a magic word for the map $R \to S$ of Definition 2.2, then the magic word isomorphism of Definition 2.2 induces an obvious shift-commuting Borel injection γ from S_W into T. The word W can be chosen so that γ has an inverse similarly defined on a subset of some $T_{\mathcal{W}'}$ in T, and then S_W can be used for the set K in Proposition 2.3. For lighter notation, we will generally use the same symbol γ for a map defined in this way and for its induced maps on functions and measures.

Properly speaking, then, almost isomorphisms relate not functions in themselves but rather equivalence classes, as follows. Given Borel functions f,g defined on subsets of S, we say f and g are *somewhere equivalent* ($f = g$ s.e.) if there exists some $S_{\mathcal{W}}$ such that f and g are defined and equal on all of $S_{\mathcal{W}}$. When we say below that two functions correspond mod s.e. under some almost isomorphism $R \to S$, $R \to T$, we mean that some representatives f, g of their s.e. equivalence classes correspond, i.e., satisfy $g = f \circ \gamma$ s.e. where γ is defined as above.

In considering the behavior of potentials and equilibrium measures under almost isomorphism, we are interested in classes of functions invariant under almost isomorphism; in particular, these functions should be described by properties which persist under restriction to smaller sets S_W. Note, if a measure does not have full support, then it will assign measure zero to some S_W. Conversely, if $f = g$ s.e., then $f = g$ ν-a.e. for all ν in $\mathcal{M}^{\text{erg}}_{\text{supp}}(S)$. Therefore we are interested in functions all of whose equilibrium measures have full support.

First we define some regularity conditions related to full support (and uniqueness) of equilibrium measures. Let S be a Markov shift and let $\mathcal{W}$ denote a set of nonempty S-words.

Definition 2.4. A function $f : S \to \mathbb{R}$ has *eventually* p*-summable variations relative to* $\mathcal{W}$ if there exists a sequence $\omega_1 \geq \omega_2 \geq \dots$ with

$$\sum_{n \geq 1} n^p \omega_n < \infty$$

such that for all x, y in S and integers $m, n \geq 0$, we have $|f(x) - f(y)| \leq \omega_n + \omega_m$ whenever

(1) $x[-m, n] = y[-m, n]$,
(2) $x[-m, -1]$ begins with a word from $\mathcal{W}$, and
(3) $x[1, n]$ ends with a word from $\mathcal{W}$.

We remark that conditions (2) and (3) imply that $m, n \geq 1$. On the other hand, if we removed these two conditions and set $\omega_n = \kappa^n$ for some $0 < \kappa < 1$, we would

[1]Dealing with ergodic invariant probability measures, we could equally require that *some* $\mathcal{W}$-word is seen infinitely often in the past and future.

be describing Hölder continuity with respect to a suitable metric – a metric "of the first type" (1.2) in [9].

Given $p \geq 0$, we define $\mathcal{E}_p(S, \mathcal{W})$ as the set of functions with eventually p-summable variations relative to $\mathcal{W}$, and set

$$\mathcal{E}_p(S) = \bigcup_{\mathcal{W}} \mathcal{E}_p(S, \mathcal{W}) \quad .$$

Like the functions with summable variations in [18], the functions in $\mathcal{E}_p(S)$ are not necessarily bounded. We let $\mathcal{E}_{p+}(S)$ denote the subset of $\mathcal{E}_p(S)$ consisting of functions depending only on future coordinates, i.e., $f(x) = f(x_0, x_1, \dots)$. For such f, the conditions of Definition 2.4 simplify (because any word can be continued to the left to a word in $\mathcal{W}$): if $x[0,n] = y[0,n]$, with $x[1,n]$ ending with a word from $\mathcal{W}$, then $|f(x) - f(y)| \leq \omega_n$.

The regularity of the potential is not sufficient to guarantee existence of an equilibrium measure (even the zero potential may fail to have an equilibrium measure [7]). For existence, we will use a notion from [18] (generalized from [7]). The *local partition function* of (S, f) at some cylinder $[W]$ associates to each positive integer n the sum

$$Z_n(S, f, W) = \sum_{\substack{S^n x = x \\ x[0,|W|-1] = W}} \exp\big(f(x) + f(Sx) + \cdots + f(S^{n-1}x)\big) \quad .$$

The pair (S, f) is said to be *positively recurrent* if for some (or equivalently, every) nonempty S-word W there exists an integer $n_0 = n_0(W)$ such that the sequence

$$\Big(Z_n(S, f, W) \exp(-nP(S, f))\Big)_{n \geq n_0}$$

is bounded away from 0 and ∞.

Definition 2.5. $\mathcal{E}(S)$ denotes the set of *relatively regular* functions on S. These are the positive recurrent Borel measurable functions f from S to $\mathbb{R} \cup \{-\infty\}$ such that $P(S, f) < \infty$ and there exists a nonempty family $\mathcal{W}$ of S-words (which may depend on f) such that $f \in \mathcal{E}_1(S, \mathcal{W}) \cup \mathcal{E}_{0+}(S, \mathcal{W})$ and every ergodic equilibrium measure of f assigns measure 1 to $S_{\mathcal{W}}$. The set of bounded relatively regular functions is denoted $\mathcal{E}^\flat(S)$.

Let us point out that in the special case that $\mathcal{W} = \mathcal{A}(S)$ (the alphabet of S), the condition in Definition 2.5 that $S_{\mathcal{W}}$ has full measure follows from the others (see [18]).

The following theorem, our main result, shows that the bounded relatively regular functions comprise one good class of potentials for almost isomorphism.

Theorem 2.6. *Suppose $R \to S$, $R \to T$ is an almost isomorphism of Markov shifts S and T. This almost isomorphism induces a mod-s.e. bijection $\gamma : \mathcal{E}^\flat(S) \to \mathcal{E}^\flat(T)$. For every $f \in \mathcal{E}^\flat(S)$, there is a unique equilibrium measure, μ_f; this measure μ_f has full support; and the almost isomorphism induces a correspondence $\gamma : \mu_f \to \mu_{(\gamma f)}$.*

We summarize Theorem 2.6 by saying that the almost isomorphism identifies the thermodynamic formalisms of the bounded relatively regular functions on S and T.

We now turn to somewhat more general results, which combine to give Theorem 2.6 as a corollary: Propositions 3.1, 4.3 and 5.2 imply Theorem 2.6.

3. Uniqueness

In this section, we prove the following uniqueness result.

Proposition 3.1. *Let S be a Markov shift and $f \in \mathcal{E}_1(S,\mathcal{W}) \cup \mathcal{E}_{0+}(S,\mathcal{W})$. Assume that f is upper bounded and that the pressure $P(S,f)$ is finite.*

Then (S,f) has at most one ergodic equilibrium measure μ giving positive measure to $\bigcup_{W\in\mathcal{W}}[W]$, and this measure must have full support.

We first note that restriction to a subsystem $S_\mathcal{W}$ does not affect the pressure.

Lemma 3.2. *Suppose $\mathcal{W}$ is a nonempty set of S-words and f is a Borel function from S to $\mathbb{R}$. Let $f_\mathcal{W}$ denote the restriction of f to $S_\mathcal{W}$. Then $P(S,f) = P(S_\mathcal{W}, f_\mathcal{W})$.*

Proof. The inequality $P(S,f) \geq P(S_\mathcal{W}, f_\mathcal{W})$ is trivial. For the reverse inequality, let μ_n be a sequence of ergodic measures of S with $\lim_n P(S,f,\mu_n) = P(S,f)$ and choose a word W in $\mathcal{W}$. Approximate μ_n by an ergodic measure ν_n such that $P(S,f,\nu_n) \geq P(S,f,\mu_n) - 1/n$, and also $\nu_n([W]) > 0$. (For example, for ν_n use a k-step Markov perturbation of the k-step Markov approximation to μ_n, for sufficiently large k.) Then $P(S_\mathcal{W}, f_\mathcal{W}) \geq \lim_n P(S,f,\nu_n) = P(S,f)$. □

Proof of 3.1. Let $f \in \mathcal{E}_1(S,\mathcal{W}) \cup \mathcal{E}_{0+}(S,\mathcal{W})$, for some collection of words $\mathcal{W}$. Assume that f is upper bounded and that $P(S,f) < \infty$.

To prove uniqueness, we assume that μ_1 and μ_2 are two ergodic equilibrium measures assigning positive measure to $\cup_{W\in\mathcal{W}}[W]$, and show that they must coincide. Choose words $w_1, w_2 \in \mathcal{W}$ such that $\mu_1([w_1]) > 0$ and $\mu_2([w_2]) > 0$. We assume that these words have the same length L, by lengthening the shorter one if necessary. For $i = 1,2$, we choose nonempty words a_i, b_i such that $\mu_i([W_i]) > 0$ for $W_i = w_i a_i w_i b_i$. We arrange it so that W_1 and W_2 have a common length N.

Let $\mathcal{G}$ be the infinite graph whose vertices are the S-words of length N, with an edge from vertex $u_0 \cdots u_{N-1}$ to vertex $u_1 \cdots u_N$ iff $u_0 u_1 \cdots u_N$ is an S-word, and we label such an edge u_0. As is well known, this graph determines a Markov shift, and the labeling defines a block code to S which is a topological conjugacy, i.e., a shift-commuting homeomorphism.

Now define another graph $\overline{\mathcal{G}}$ as follows. $\overline{\mathcal{G}}$ will include two distinguished vertices $\overline{v_1}$ and $\overline{v_2}$. For each path $p = p_1 \cdots p_k$ in $\mathcal{G}$ with initial vertex W_i and terminal vertex W_j, and no intermediate vertices equal to W_i or W_j, $\overline{\mathcal{G}}$ contains a path $\overline{p}$ of equal length from $\overline{v_i}$ to $\overline{v_j}$. These paths contain all edges of $\overline{\mathcal{G}}$ and overlap only at their initial and terminal vertices in $\{\overline{v_1}, \overline{v_2}\}$. The S-word of length k labeling the edges $p_1, p_2, \ldots, p_k$ is used to likewise label the k edges $\overline{p}_1, \overline{p}_2, \ldots, \overline{p}_k$. Let R be the Markov shift defined from the resulting graph and let $\varphi : R \to S$ be the injective block code defined by the edge labeling. The image of φ has measure one for both μ_1 and μ_2. Let $\overline{f}$ denote the function $f \circ S^L \circ \varphi$.

Let $\overline{\mathcal{A}}$ denote the alphabet of R. We claim that

$$\overline{f} \in \mathcal{E}_1(R, \overline{\mathcal{A}}) \cup \mathcal{E}_{0+}(R, \overline{\mathcal{A}}) \tag{3.3}$$

(contrarily to f on S). Suppose $\overline{x}, \overline{y}$ are sequences in R with $\overline{x}[-m,n] = \overline{y}[-m,n]$ for some integers $m, n \geq 0$ which we may and do assume to be taken maximum (the case where $\overline{x}$ and $\overline{y}$ agree in all nonnegative coordinates or all nonpositive coordinates is similar and is left to the reader). The initial vertex of $\overline{x}_{-m}$ and the terminal vertex of $\overline{x}_n$ are contained in $\{\overline{v}_1, \overline{v}_2\}$. Let $x = \varphi(\overline{x})$ and $y = \varphi(\overline{y})$. Then for some $i, j \in \{1, 2\}$,

$$x[-m, -m+L-1] = y[-m, -m+L-1] = w_i$$

and since $x[n, n+N-1] = w_j a_j w_j b_j$,

$$x[n+N-|w_j b_j|, n+N-|b_j|-1] = y[n+N-|w_j b_j|, n+N-|b_j|-1] = w_j.$$

Remark that

$$\begin{aligned} (-m+L-1)-L &\leq -1-m\ , \quad \text{and} \\ (n+N-|w_j b_j|)-L &= n+|a_j| \geq n+1 \end{aligned}$$

so $f \in \mathcal{E}_p(S, \{w_1, w_2\})$ gives:

$$|\overline{f}(\overline{x}) - \overline{f}(\overline{y})| = |f(S^L x) - f(S^L y)| \leq \omega_{m+L} + \omega_{n+M}$$

where $M = N - |b_j| - L - 1 \geq 0$, and the sequence (ω_n) comes from Definition 2.4 for f. Now define $\overline{\omega}_n = \max\{\omega_{n+L}, \omega_{n+M}\}$, so

$$|\overline{f}(\overline{x}) - \overline{f}(\overline{y})| \leq \overline{\omega}_m + \overline{\omega}_n$$

and

$$\sum_{n=1}^{\infty} n^p \overline{\omega}_n \leq \sum_{n=1}^{\infty} (n+L)^p \omega_{n+L} + (n+M)^p \omega_{n+M} < \infty\ .$$

Moreover if f depends only on future coordinates, then this is obviously the case for $\overline{f}$. Therefore the claim (3.3) is proved (in particular $\overline{f}$ has bounded oscillation on cylinders $[a]$, a a vertex of $\overline{\mathcal{G}}$, contrarily to f).

Next, in the case where f also depends on past coordinates, we make the following standard replacement to obtain a function depending only on future coordinates. Following [1], let $h : R \to \mathbb{R}$ be given by :

$$h(x) = \sum_{k \geq 0} \overline{f}(R^k x) - \overline{f}(R^k x^-) \qquad \forall x \in R$$

where $x^- \in R$ with $x_i^- = x_i$ for $i \geq 0$ and $(x_i^-)_{i \leq 0}$ depending only on x_0. The above series is bounded by $\sum_{k \geq 0} \omega_k$, and converges. One easily checks that h is bounded and has summable variations. One also sees that $\overline{f} + h \circ R - h$ depends only on $(x_i)_{i \geq 0}$.

If f depends only on future coordinates, let $h = 0$. Then in both cases $\overline{f} + h \circ R - h$ is upper bounded, has summable variations and depends only on future coordinates. We can apply the main theorem of [4] and see that $(R, \overline{f})$ has at most one ergodic equilibrium measure, and that this measure has full support when it exists.

To bring back this result to (S, f) first notice that $\mu(\overline{f} + h \circ R - h) = \mu(f \circ \varphi)$ for all R-invariant probability measures. We claim further that $P(R, \overline{f}) = P(S, f)$. This follows from Lemma 3.2 because φ is a continuous bijection, with Borel measurable inverse function, and the image of φ is $S_{\{W_1, W_2\}}$. Thus μ_1 and μ_2 must correspond under φ to the unique equilibrium measure of $(R, \overline{f})$. Hence $\mu_1 = \mu_2$.

Because φ is continuous and has dense image, it also follows that $\mu_1 = \mu_2$ has full support in S. $\square$

4. Existence

The issue of existence in our context is more complicated. As in [18, 23, 25], we will use the concept of a *weak equilibrium measure* for (S, f): an invariant probability measure μ for which $P(S, f, \mu)$ is not necessarily defined, but which for some measurable function h satisfies the following:

$$-\sum_{a\in\mathcal{A}} 1_{[a]} \log E_\mu(1_{[a]}|S^{-1}\mathcal{B}) + f + h \circ S - h \in L^1(\mu)$$

and

$$\int \left(-\sum_{a\in\mathcal{A}} 1_{[a]} \log E_\mu(1_{[a]}|S^{-1}\mathcal{B}) + f + h \circ S - h \right) d\mu = P(S, f), \tag{4.1}$$

where $\mathcal{A}$ is the set of states of S, $\mathcal{B}$ is the σ-algebra of Borel measurable subsets, $E_\mu(\cdot|\cdot)$ is the conditional expectation. (The function h was assumed to be locally Hölder-continuous in [18], but not here.)

Remark 4.2. It is possible for a positive recurrent Hölder continuous potential to have a weak equilibrium measure when there is no equilibrium measure [18, Sec.7]. When S has finite topological entropy and f is bounded, then a weak equilibrium measure for (S, f) must be an equilibrium measure. However there exist upper-bounded potentials which are positive recurrent, have finite pressure and summable variations, define a weak equilibrium measure which has finite entropy, and, yet, this measure is not a (strong) equilibrium measure.

O. Sarig pointed out to us the following example of such a potential. Fix $1/2 < \alpha < 1$ and take the interval map $T : [0, 1] \to [0, 1]$ defined by $T(x) = x(1 + 2^\alpha x^\alpha)$ for $x < 1/2$ and $T(x) = 2x - 1$ otherwise [13]. Let $\varphi(x) = -\log|T'| - \alpha \log(x/Tx) \le 0$. Using the partition of $[0, 1] \setminus \bigcup_{k\ge0}(T^{-k}(1/2) \cap [0, 1/2))$ into its connected components, one obtains from (T, φ) a Markov shift with a potential which gives the required example (use [19, 20] for the summable variations and [21] for estimating the density of the weak equilibrium measure). We do not know if one can find an example where additionally the potential is Hölder-continuous.

We now state our main existence result.

Proposition 4.3. *Let S be a Markov shift and $f \in \mathcal{E}_1(S, \mathcal{W}) \cup \mathcal{E}_{0+}(S, \mathcal{W})$. Assume that f is upper bounded and the pressure $P(S, f)$ is finite.*

If (S, f) is positive recurrent, then there exists a weak equilibrium measure. Conversely, if there exists an equilibrium measure μ with $\mu S_\mathcal{W} = 1$, then (S, f) is positive recurrent.

If f is bounded and $\mu S_\mathcal{W} = 1$ whenever μ is an equilibrium measure for f, then (S, f) is positive recurrent if and only if f has an equilibrium measure.

Remark 4.4. Propositions 3.1 and 4.3 are known (in the case $S_\mathcal{W} = S$) if f is uniformly locally constant [9] or if it is uniformly Hölder continuous or more generally has summable variations and depends only on the future [18]. We note that Gibbs measures (see [14]) are a rather different issue [20].

Remark 4.5. We do not know if the existence of a weak equilibrium measure implies that a potential in $\mathcal{E}_1(S, \mathcal{W}) \cup \mathcal{E}_{0+}(S, \mathcal{W})$ is positive recurrent.

Remark 4.6. The paper [5] gives more general sufficient conditions for existence of equilibrium measures of continuous functions on a Markov shift. We have not exploited those conditions here because we do not know when the equilibrium measures produced in [5] have full support, and we do not know whether the Z-recurrence condition in [5] must be preserved under passage to a smaller system S_W.

Remark 4.7. $f \in \mathcal{E}_p(S, \mathcal{W})$ is arbitrary on the subshift S_* of S obtained by excluding all words in $\mathcal{W}$. Hence, the exclusion in Definition 2.5 of equilibrium measures living on S_* is necessary for drawing conclusions about equilibrium measures on S. Also, in the definition we could equivalently require f to be positive recurrent on $S_{\mathcal{W}}$ rather than on S, since given f on $S_{\mathcal{W}}$ we could extend f to S without changing the pressure or set of equilibrium measures

Proof of 4.3. We continue with the notation and constructions used in the proof above for Prop. 3.1. Let $(R, \overline{f})$ be constructed as in that proof, using as $w_1 = w_2$ some word from $\mathcal{W}$ to be specified.

Assume first that (S, f) is positive recurrent. Take $w_1 = w_2 \in \mathcal{W}$ arbitrarily. Positive recurrence of $(R, \overline{f})$ follows from $P(R, \overline{f}) = P(S, f)$ and the coincidence of local partition functions for all $n \geq 1$,

$$Z_n(S, f, W_1) = Z_n(R, \overline{f}, \overline{v}_1)$$

where W_1 is a word in $\mathcal{W}$ and $\overline{v}_1$ is an element of the alphabet of R, both defined as in the proof for Prop. 3.1. Then $(R, \overline{f})$ has a weak equilibrium measure by Theorem 7 of [18]. $\overline{\varphi}$ being one-to-one, this gives a weak equilibrium measure for (S, f).

Conversely, assume now that (S, f) has an equilibrium measure μ with $\mu(S_W) = 1$. By assumption, we can choose some word $w_1 = w_2 \in \mathcal{W}$ with $\mu([w_1]) > 0$. μ can then be pulled back to R by $\overline{\varphi}$. As $P(R, \overline{f}) = P(S, f)$ the pullback of μ is an equilibrium measure for $(R, \overline{f})$. This implies the positive recurrence of $(R, \overline{f})$ according to [18] and [4] ([4] says that any equilibrium measure satisfies Sarig's Ruelle-Perron-Frobenius theorem; [18] says that existence of such an invariant measure implies positive recurrence). The positive recurrence of (S, f) follows as above.

It remains to show that when f is bounded, a weak equilibrium measure for f must be an equilibrium measure. This is an immediate consequence of Lemma 4.8 below. □

Lemma 4.8. *Let $f \in \mathcal{E}_0(S, \mathcal{W})$ be upper-bounded with $P(S, f) < \infty$. If μ is a weak equilibrium measure for (S, f) such that $f \in L^1(\mu)$ and $\mu\left(\bigcup_{W \in \mathcal{W}} [W]\right) > 0$ then μ is an equilibrium measure.*

Proof of Lemma 4.8. Let h be the measurable function given by the assumption that μ is a weak equilibrium measure. Recall our notation $S_n f = f + f \circ S + \cdots + f \circ S^{n-1}$.

The first point to see is that $\int h - h \circ S \, d\mu = 0$. We follow Ledrappier [11] by observing that $h - h \circ S$ is integrable as the difference of two integrable functions, namely f and $f + (h - h \circ S)$, and therefore Birkhoff's ergodic theorem gives

$$\int h - h \circ S \, d\mu = \lim_{n \to \infty} \frac{1}{n} S_n(h - h \circ S) = \lim_{n \to \infty} \frac{1}{n}(h - h \circ S^n) \quad \text{a.e.}$$

Recurrence implies that 0 is an accumulation point for the sequence $\frac{1}{n}(h - h \circ S^n)$ a.e. Therefore the above limit is zero, proving the first point and therefore

$$\int f\,d\mu = \int f + h - h \circ S\,d\mu \ . \tag{4.9}$$

We have the following properties:

- $P(S,f) < \infty$;
- $f \in L^1(\mu)$;
- because of the eventual summable variation property, there exists a cylinder $[W_*]$ with $\mu([W_*]) > 0$ such that for all $x, y \in [W_*]$, $n \geq 0$ with $S^n(x), S^n(y) \in [W_*]$, $|S_n f(x) - S_n f(y)| <$ const.

The last property is proved as in the proof of Proposition 3.1 (as there, we replace f by $f \circ S^L$). We claim that

$$\int -\sum_{a \in \mathcal{A}} 1_{[a]} \log E_\mu(1_{[a]}|T^{-1}\mathcal{B})\,d\mu \ = \ h(S,\mu) \tag{4.10}$$

despite the fact that the partition $\{[a] : a \in \mathcal{A}\}$ may have infinite entropy. This is proved in [4, p. 1389], assuming the three properties listed above. Because μ is a weak equilibrium measure, the facts (4.9) and (4.10) together show that μ is indeed an equilibrium measure. □

5. Correspondence

Definition 5.1. $\mathcal{E}^{ub}(S)$ is the class of upper bounded functions f with $P(S,f) < \infty$ such that there exists $\mathcal{W}$ (allowed to depend on f) such that $f \in \mathcal{E}_1(S,\mathcal{W}) \cup \mathcal{E}_{0+}(S,\mathcal{W})$ and for any equilibrium measure μ of f, $\mu[W] > 0$ for some $W \in \mathcal{W}$.

Proposition 5.2. *Suppose $R \to S$, $R \to T$ is an almost isomorphism of Markov shifts S and T. This almost isomorphism induces mod-s.e. bijections $\Gamma : \mathcal{E}_p(S) \to \mathcal{E}_p(T)$ and $\Gamma : \mathcal{E}^{ub}(S) \to \mathcal{E}^{ub}(T)$. Moreover, for $f \in \mathcal{E}^{ub}(S)$ we have:*

- *(S,f) is positive recurrent if and only if $(T,\Gamma(f))$ is positive recurrent;*
- *an invariant probability measure μ of S is an equilibrium measure, resp. fully supported weak equilibrium measure, for (S,f) if and only if $\gamma\mu$ is an equilibrium measure, resp. fully supported weak equilibrium measure, for $(T,\Gamma(f))$.*

Remark 5.3. We do not know whether a weak equilibrium measure for a Hölder continuous potential is always fully supported.

Proof. Let $\gamma : K \subset S \to K' \subset T$ be the bijection given by the almost isomorphism according to Proposition 2.3. We first prove that if $f \in \mathcal{E}_p(S,\mathcal{W})$ then for some collection $\mathcal{V}$ of T-words,

$$g = f \circ \gamma^{-1} : K' \to \mathbb{R} \in \mathcal{E}_p(T,\mathcal{V}) \ . \tag{5.4}$$

(If one likes, one can extend g to the whole of T.)

Let $W_0 \in \mathcal{W}$. Let V_1 be some magic T-word for the map $R \to T$ such that K' contains all points in which V_1 occurs infinitely often in the future and the past. Choose a T-word V_2 and an integer $J \geq 0$ such that for all $y \in K'$ and $i \in \mathbb{Z}$,

$$\begin{aligned} \text{if} \ \ & y[i, i + |V_1V_2V_1| - 1] = V_1V_2V_1 \ , \\ \text{then} \ \ & (\gamma^{-1}y)[i + J, i + J + |W_0| - 1] = W_0 \ . \end{aligned}$$

We may also assume $J \leq |V_1V_2V_1| - |W_0|$. Let V be the T-word $V_1V_2V_1$. We claim that $g \in \mathcal{E}_p(T, \mathcal{V})$ with $\mathcal{V} = \{V\}$.

Indeed let $x, y \in K'$ and the integers m, n be such that $x[-m, n] = y[-m, n]$ begins and ends with the word V. Then $x = \gamma(u)$ and $y = \gamma(v)$, where

$$u[-m + J, n - |V| + J + |W_0| + 1] \ = \ v[-m + J, n - |V| + J + |W_0| + 1]$$

and this word begins and ends with W_0. Therefore

$$|g(x) - g(y)| \ = \ |f(u) - f(v)| \ \leq \ \omega_{m-J} + \omega_{n-|V|+J}$$

which implies (5.4) by an argument similar to the proof of (3.3).

We next check that Γ is well-defined on the level of somewhere equivalent classes. Let $f_1, f_2 \in \mathcal{E}_p(S)$ be somewhere equivalent, i.e., for some word w, $f_1 = f_2$ over S_w. By enlarging the word W_0 in the above construction to a word W which contains w, we obtain that $T_V \subset \gamma(S_W)$. This implies that $\Gamma(f_1) = \Gamma(f_2)$ s.e. Because $\gamma : K \to K'$ is a bijection, the induced map Γ on somewhere equivalent classes is also a bijection.

We now check that for $f \in \mathcal{E}^{u\flat}(S)$, there exists $g \in \mathcal{E}^{u\flat}(T)$ somewhere equivalent to $f \circ \gamma^{-1}$. We define g to be $f \circ \gamma^{-1}$ on K'. Let T_* denote the restriction of T to the complement K'_* of K'; choose an upper bounded function g_* on K'_* such that $P(T_*, g_*) < P(T, g)$; and on K'_*, define $g = g_*$. Now any equilibrium measure for (g, T) must be supported on K'. Using Lemma 3.2, we have

$$P(T, g) = P(T|K', g) = P(S|K, f|K) = P(S, f) < \infty \ .$$

This proves that $\Gamma(\mathcal{E}^{u\flat}(S)) \subset \mathcal{E}^{u\flat}(T)$ modulo s.e. It follows that $\Gamma : \mathcal{E}^{u\flat}(S) \to \mathcal{E}^{u\flat}(T)$ is a bijection modulo s.e.

Note that Proposition 3.1 says that if μ is an equilibrium measure for $f \in \mathcal{E}^{u\flat}(S)$, it is ergodic and has full support. By Proposition 2.3, $(T, \gamma\mu)$ is isomorphic to (S, μ) as soon as μ is fully supported. Thus we see that the equilibrium measure μ of (S, f) corresponds to a measure $\gamma\mu$ with the same pressure for $(T, \Gamma(f))$. Because $P(S, f) = P(T, \Gamma(f))$, $\gamma\mu$ is the equilibrium measure for $(T, \Gamma(f))$ as claimed.

The same reasoning applies to fully supported weak equilibrium measures. $\square$

References

[1] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math. **470** (1975), Springer-Verlag.

[2] M. Boyle, J. Buzzi and R. Gómez, *Almost isomorphism for countable state Markov shifts*, J. für Ang. und Reine Math., to appear.

[3] J. Buzzi, *Subshifts of quasi-finite type,* Invent. Math. **159** (2005), no. 2, 369–406.

[4] J. Buzzi, O. Sarig, *Uniqueness Of Equilibrium Measures For Countable Markov Shifts And Multidimensional Piecewise Expanding Maps*, Ergod. Th. & Dynam. Syst. **23** (2003), 1383–1400.

[5] D. Fiebig, U.-F. Fiebig and M. Yuri, *Pressure and equilibrium states for countable state Markov shifts*, Israel J. Math. 131 (2002), 221–257.

[6] U.-R. Fiebig, *Symbolic dynamics and locally compact Markov shifts*, 1996. Habilitationsschrift, U. Heidelberg.

[7] B. M. Gurevich, *Shift entropy and Markov measures in the path space of a denumerable graph* (Russian), Dokl. Akad. Nauk SSSR **187** (1969), 715–718; English translation: Soviet Math. Dolk. **10**, 4, 911–915.

[8] B. M. Gurevich, *Stably recurrent nonnegative matrices* (Russian), Uspekhi Mat. Nauk **51** (1996), no. 3(309), 195–196; translation in Russian Math. Surveys **51** (1996), no. 3, 551–552.

[9] B. M. Gurevich and S. Savchenko, *Thermodynamic formalism for symbolic Markov chains with a countable number of states* (Russian), Uspekhi Mat. Nauk **53** (1998), 3–106; translation in Russian Math. Surveys **53** (1998), 245–344.
[10] B.P. Kitchens, *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts*, Springer-Verlag (1998).
[11] F. Ledrappier, *Principe variationnel et systèmes dynamiques symboliques*, Z. Wahrsch. Verw. Gebiete **30** (1974), 185–202.
[12] D. Lind and B. Marcus, *An introduction to Symbolic Dynamics and Coding*, Cambridge University Press (1995).
[13] C. Liverani, B. Saussol, S. Vaienti, *Conformal measure and decay of correlation for covering weighted systems,* Ergodic Theory Dynam. Systems **18** (1998), 1399–1420.
[14] R. D. Mauldin and M. Urbański, *Gibbs states on the symbolic space over an infinite alphabet,* Israel J. Math. **125** (2001), 93–130.
[15] K. Petersen, *Ergodic theory*, corrected reprint of the 1983 original. Cambridge University Press, Cambridge, 1989.
[16] S. Ruette, *On the Vere-Jones classification and existence of maximal measures for countable topological Markov chains*, Pacific J. Math. **209** (2003), 365–380.
[17] I. Salama, *On the recurrence of countable topological Markov chains*, in Symbolic dynamics and its applications (New Haven, CT, 1991), Contemp. Math., **135** (1992), 349–360, Amer. Math. Soc., Providence, RI,
[18] O. Sarig, *Thermodynamic formalism for countable Markov shifts*, Ergodic Theory Dynam. Systems **19** (1999), no. 6, 1565–1593.
[19] O. Sarig, *Subexponential decay of correlations,* Invent. Math. **150** (2002), 629-653.
[20] O. Sarig, *Existence of Gibbs measures for countable Markov shifts*, Proc. Amer. Math. Soc. **131** (2003), no. 6, 1751–1758
[21] M. Thaler, *Estimates of the invariant densities of endomorphisms with indifferent fixed points,* Israel J. Math. **37** (1980), 303–314.
[22] D. Vere-Jones, *Geometric ergodicity in denumerable Markov chains*, Quart. J. Math. Oxford (2) **13** (1962), 7–28.
[23] P. Walters, *Ruelle's operator theorem and g-measures*, Trans. Amer. Math. Soc. **214** (1975), 375–387.
[24] P. Walters, *An introduction to ergodic theory*, Springer-Verlag (1982).
[25] M. Yuri, *On the convergence to equilibrium states of certain non-hyperbolic systems*, Ergod. Th. & Dynam. Syst. **17** (1997), 977-1000.

Mike Boyle, Department of Mathematics, University of Maryland, College Park, MD 20742-4015, U.S.A.
E-mail address: `mmb@math.umd.edu`
URL: `www.math.umd.edu/∼mmb`

Jerome Buzzi, Centre de Mathématiques, Ecole polytechnique, 91128 Palaiseau Cedex, France
E-mail address: `buzzi@math.polytechnique.fr`
URL: `www.jeromebuzzi.com`

Ricardo Gómez, Instituto de Matematicas Area de la Investigacion Cientifica, Circuito Exterior, Ciudad Universitaria, DF 04510, Mexico
E-mail address: `rgomez@math.unam.mx`
URL: `www.math.unam.mx/∼rgomez`